

THE PARAMETRIZED FAMILY OF METRIC MAHLER MEASURES

CHARLES L. SAMUELS

ABSTRACT. Let $M(\alpha)$ denote the (logarithmic) Mahler measure of the algebraic number α . Dubickas and Smyth, and later Fili and the author, examined metric versions of M . The author generalized these constructions in order to associate, to each point in $t \in (0, \infty]$, a metric version M_t of the Mahler measure, each having a triangle inequality of a different strength. We further examine the functions M_t , using them to present an equivalent form of Lehmer's conjecture. We show that the function $t \mapsto M_t(\alpha)^t$ is constructed piecewise from certain sums of exponential functions. We pose a conjecture that, if true, enables us to graph $t \mapsto M_t(\alpha)$ for rational α .

1. INTRODUCTION

Let f be a polynomial with complex coefficients given by

$$f(z) = a \cdot \prod_{n=1}^N (z - \alpha_n).$$

We define the (*logarithmic*) *Mahler measure* M of f by

$$M(f) = \log |a| + \sum_{n=1}^N \log^+ |\alpha_n|.$$

If α is a non-zero algebraic number, we define the (*logarithmic*) *Mahler measure* $M(\alpha)$ of α to be the Mahler measure of the minimal polynomial of α over \mathbb{Z} .

It is a consequence of a theorem of Kronecker that $M(\alpha) = 0$ if and only if α is a root of unity. In a famous 1933 paper, D.H. Lehmer [5] asked whether there exists a constant $c > 0$ such that $M(\alpha) \geq c$ in all other cases. He could find no algebraic number with Mahler measure smaller than that of

$$\ell(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1,$$

which is approximately $0.16 \dots$. Although the best known general lower bound is

$$M(\alpha) \gg \left(\frac{\log \log \deg \alpha}{\log \deg \alpha} \right)^3,$$

due to Dobrowolski [2], uniform lower bounds have been established in many special cases (see [1, 13, 14], for instance). Furthermore, numerical evidence provided, for example, in [6–9] suggests there does, in fact, exist such a constant c . This leads to the following conjecture, which we will now call Lehmer's conjecture.

2000 *Mathematics Subject Classification*. 11R04, 11R09 (Primary), 30D20, 54A10 (Secondary).
Key words and phrases. Weil height, Mahler measure, metric Mahler measure, Lehmer's problem.
 This research was supported in part by NSERC of Canada.

Conjecture 1.1 (Lehmer's conjecture). *There exists a real number $c > 0$ such that if $\alpha \in \overline{\mathbb{Q}}^\times$ is not a root of unity then $M(\alpha) \geq c$.*

Dubickas and Smyth [3], and later Fili and the author [4], examined metric and ultrametric versions of the Mahler measure on $\overline{\mathbb{Q}}$, respectively. In [12], we noted that these constructions arise from the following more general principle.

Let G be an abelian group (written multiplicatively) with identity e . We say that $\phi : G \rightarrow [0, \infty)$ is a (*logarithmic*) *height* on G if the following two conditions are satisfied.

- (i) $\phi(e) = 0$,
- (ii) $\phi(\alpha) = \phi(\alpha^{-1})$ for all $\alpha \in G$.

If ψ is another height on G , we follow the conventional notation that

$$\phi = \psi \quad \text{or} \quad \phi \leq \psi$$

when $\phi(\alpha) = \psi(\alpha)$ or $\phi(\alpha) \leq \psi(\alpha)$ for all $\alpha \in G$, respectively. We write

$$Z(\phi) = \{\alpha \in G : \phi(\alpha) = 0\}$$

to denote the *zero set* of ϕ .

If t is a positive real number then we say that ϕ has the *t-triangle inequality* if

$$(1.1) \quad \phi(\alpha\beta)^t \leq \phi(\alpha)^t + \phi(\beta)^t$$

for all $\alpha, \beta \in G$. We say that ϕ has the *∞ -triangle inequality* if

$$(1.2) \quad \phi(\alpha\beta) \leq \max\{\phi(\alpha), \phi(\beta)\}$$

for all $\alpha, \beta \in G$. We observe that the 1-triangle inequality is simply the classical triangle inequality while the ∞ -triangle inequality is the strong triangle inequality. A height ϕ satisfying (1.1) or (1.2) is called a *t-metric height* or *∞ -metric height*, respectively. It is noted in [12] that such heights have the following properties.

- (i) $Z(\phi)$ is a subgroup of G .
- (ii) ϕ is well-defined on the quotient $G/Z(\phi)$.
- (iii) If $t \geq 1$, then the map $(\alpha, \beta) \mapsto \phi(\alpha\beta^{-1})$ defines a metric on $G/Z(\phi)$.

If ϕ is a height which is not necessarily a *t-metric height*, then we may construct a natural *t-metric version* of ϕ . For simplicity, we will now write

$$\mathcal{X}(G) = \{(\alpha_1, \alpha_2, \dots) : \alpha_n \in G \text{ and } \alpha_n = e \text{ for all but finitely many } n\}.$$

If \mathbb{R} denotes the group of real numbers under addition, $\mathbf{x} = (x_1, x_2, \dots) \in \mathcal{X}(\mathbb{R})$, and t is any positive real number, we define

$$(1.3) \quad \|\mathbf{x}\|_t = \left(\sum_{n=1}^{\infty} |x_n|^t \right)^{1/t} \quad \text{and} \quad \|\mathbf{x}\|_{\infty} = \max_{1 \leq n} \{|x_n|\}.$$

In the case where $t \geq 1$, we know that $\|\mathbf{x}\|_t$ is the L^t norm of \mathbf{x} . If $t < 1$, then (1.3) does not define a norm on $\mathcal{X}(\mathbb{R})$, but we continue to use the same notation for the sake of consistency. Let $\tau : \mathcal{X}(G) \rightarrow G$ be defined by

$$\tau(\alpha_1, \alpha_2, \dots) = \prod_{n=1}^{\infty} \alpha_n$$

and note that τ is a group homomorphism. The *t-metric version* of ϕ is given by

$$\phi_t(\alpha) = \inf \{ \|\phi(\alpha_1), \phi(\alpha_2), \dots\|_t : (\alpha_1, \alpha_2, \dots) \in \tau^{-1}(\alpha) \}$$

so that the infimum is taken over all ways of writing α as a product of elements in G . It is immediately clear that if ψ is another height on G with $\phi \geq \psi$, then $\phi_t \geq \psi_t$ for all t . The results of [12] establish the following additional observations.

- (i) ϕ_t is a t -metric height on G with $\phi_t \leq \phi$.
- (ii) If ψ is an t -metric height with $\psi \leq \phi$ then $\psi \leq \phi_t$.
- (iii) $\phi = \phi_t$ if and only if ϕ is an t -metric height.
- (iv) If $s \in (0, t]$ then $\phi_s \geq \phi_t$.

It is well-known that the Mahler measure M is a height on $\overline{\mathbb{Q}}^\times$ with $Z(M)$ equal to the set of roots of unity. It follows from the results of [3] and [4] that $Z(M_t) = Z(M)$ for all $t \in (0, \infty]$. Among other things, it is noted that M_1 and M_∞ induce the discrete topology on

$$V = \overline{\mathbb{Q}}^\times / Z(M)$$

if and only if Lehmer's conjecture is true. It turns out that we have something stronger.

Theorem 1.2. *Lehmer's conjecture is true if and only if there exists $t \in [1, \infty)$ such that M_t and M_∞ induce the same topology on V .*

Our goal for the remainder of this article is to examine the functions $t \mapsto M_t(\alpha)$ for a fixed algebraic number α . For simplicity, we define $\mu_\alpha : (0, \infty] \rightarrow [0, \infty)$ by

$$\mu_\alpha(t) = M_t(\alpha).$$

It is clear from our earlier remarks that μ_α is decreasing, bounded above by $M(\alpha)$, and $\mu_\alpha(t)$ tends to $M_\infty(\alpha)$ as $t \rightarrow \infty$. The results of [12] give some additional properties of μ_α , namely

- (i) μ_α is continuous on $(0, \infty)$,
- (ii) μ_α is constant in a neighborhood of 0, and
- (iii) The infimum in the definition of $\mu_\alpha(t)$ is always attained.

This final observation suggests the following direction of study. While the set

$$A_\alpha(t) = \{\mathbf{x} \in \mathcal{X}(\mathbb{R}) : \mu_\alpha(t) = \|\mathbf{x}\|_t\}$$

is always non-empty, it is possible that $A_\alpha(t_1) \cap A_\alpha(t_2)$ is empty for different points t_1 and t_2 . This suggests that there are points $t \in (0, \infty)$ such that the point \mathbf{x} where the infimum is attained must change. We call these points α -exceptional and capture this concept rigorously in the following way.

A set $I \subseteq (0, \infty]$ is called α -uniform if there exists a point $\mathbf{x} \in \mathcal{X}(\mathbb{R})$ such that

$$\mu_\alpha(t) = \|\mathbf{x}\|_t$$

for all $t \in I$. A point $s \in (0, \infty]$ is called α -standard if there exists an α -uniform open neighborhood of s . If s is not α -standard, then we say that s is α -exceptional. Our first result shows that the set of α -exceptional points is rather sparse.

Theorem 1.3. *If α is a non-zero algebraic number and T is a positive real number, then there are only finitely many α -exceptional points in $(0, T)$.*

It is an open question to determine whether there are only finitely many α -exceptional points in all of $(0, \infty)$. The proof of Theorem 1.3 relies on an upper bound, depending on both α and T , on the number of terms that may appear in any factorization of α . It appears that we cannot remove the dependency on T to

establish the finiteness of the set of α -exceptional points. Nonetheless, we know of no example of an algebraic number α having infinitely many α -exceptional points.

Conceptually, the α -exceptional points represent values of t at which the infimum attaining point \mathbf{x} must change. Our next Theorem shows that the intervals between the α -exceptional points contain no such changes.

Theorem 1.4. *Suppose that $0 < a < b < \infty$. Then $[a, b]$ is α -uniform if and only if every point in (a, b) is α -standard. Moreover, $(0, a]$ is α -uniform if and only if every point in $(0, a)$ is α -standard.*

We now apply Theorems 1.3 and 1.4 to show that μ_α may be constructed piecewise from functions of the form $t \mapsto \|\mathbf{x}\|_t$. The pieces are divided precisely by the α -exceptional points.

Corollary 1.5. *Let α be a non-zero algebraic number and T a positive real number. There exists a finite collection of non-overlapping intervals \mathcal{I} , each closed in $(0, T]$, such that*

- (i) *Each interval in \mathcal{I} is α -uniform,*
- (ii) *$(0, T] = \cup_{I \in \mathcal{I}} I$, and*
- (iii) *If $t \in (0, T)$ then t is α -exceptional if and only if there exist distinct intervals $I_1, I_2 \in \mathcal{I}$ such that $t \in I_1 \cap I_2$.*

We now wish to establish a connection between the α -standard points and the differentiability of μ_α . Although it is clear that μ_α is infinitely differentiable at all α -standard points, it is not obvious what happens at α -exceptional points. Our next theorem gives some additional insight.

Theorem 1.6. *Let α be an algebraic number and $s \in (0, \infty)$. Then s is α -standard if and only if μ_α is infinitely differentiable at s .*

2. A CONJECTURE ON THE INFIMUM IN $M_t(\alpha)$ AND SOME APPLICATIONS

For this section, we restrict our attention to the case that α is rational. In this simpler setting, we may be able to give a more thorough description of μ_α .

Recall that Theorem 3.2 shows the infimum in the definition of $M_t(\alpha)$ to be attained. Moreover, in the case that α is rational, this infimum must be attained by a point $(\alpha_1, \dots, \alpha_N)$ where each α_n is a surd. However, we are unable to construct an example where the infimum is not attained by a point having only rational coordinates. This leads to the following conjecture.

Conjecture 2.1. *Suppose α is a rational number and $t \in (0, \infty]$. Then there exist rational points $\alpha_1, \dots, \alpha_N$ such that*

$$M_t(\alpha)^t = \sum_{n=1}^N M(\alpha_n)^t.$$

In view of the results of [3] and [4], Conjecture 2.1 is true for the cases $t \leq 1$ and $t = \infty$. In fact, in each case, a specific representation can be given that attains the infimum in $M_t(\alpha)$. Unfortunately, the proofs seem to be genuinely different and cannot be modified to include the intermediate values of t .

If Conjecture 2.1 is true, then we may often explicitly graph $\mu_\alpha(t)$. Our procedure relies on the following observation.

Theorem 2.2. *Suppose that r and s are relatively prime positive integers. If Conjecture 2.1 holds, then there exist positive integers $r_1, \dots, r_N, s_1, \dots, s_N$ such that*

$$M_t\left(\frac{r}{s}\right)^t = \sum_{n=1}^N M\left(\frac{r_n}{s_n}\right)^t$$

and

$$r = \prod_{n=1}^N r_n \quad \text{and} \quad s = \prod_{n=1}^N s_n.$$

The first statement of Theorem 2.2 is simply a rephrasing of Conjecture 2.1. The real content of the result occurs in the second statement, which shows that we need only consider all possible factorizations of the numerator and denominator. This allows us to determine $M_t(\alpha)$ with a finite search. The case where $\alpha \in \mathbb{Z}$ is particularly straightforward.

Theorem 2.3. *Suppose that α is a positive integer and write*

$$\alpha = \prod_{n=1}^N p_n$$

where p_n are not necessarily distinct primes. If Conjecture 2.1 holds then

$$M_t(\alpha)^t = \begin{cases} (\log \alpha)^t & \text{if } t \leq 1 \\ \sum_{n=1}^N (\log p_n)^t & \text{if } t \geq 1. \end{cases}$$

Theorem 2.3 shows, in particular, that under Conjecture 2.1, an integer has no exceptional points except possibly at 1. An integer has an exceptional point at 1 if and only if that integer is composite.

It is natural to ask whether a result analogous to Theorem 2.3 holds for any rational number α . Although we always have that $M_t(\alpha) = M(\alpha)$ for $t \leq 1$, the situation seems to be more complicated for larger values of t . We continue to assume Conjecture 2.1 in the remarks that follow.

Consider, for example, $\alpha = 7/30$. In the left column of Table 1, we give all possible representations of $7/30$ that satisfy the conclusion of Theorem 2.2. In the right column, we write their corresponding (non-logarithmic) Mahler measures.

We obtain immediately a natural partial ordering on the N -tuples (a_1, \dots, a_N) appearing in the right column of Table 1. We say that $(a_1, \dots, a_N) \leq (b_1, \dots, b_M)$ if

$$\|(a_1, \dots, a_N)\|_t \leq \|(b_1, \dots, b_M)\|_t$$

for all $t > 0$. For example, we note that $(2, 3, 7) \leq (2, 5, 7)$. On the other hand, the L^t norms of (30) and $(7, 15)$ cross when

$$(\log 30)^t = (\log 7)^t + (\log 15)^t$$

so that these elements are not comparable. An N -tuple (a_1, \dots, a_N) is called *minimal* if there does not exist another M -tuple (b_1, \dots, b_M) in right column of Table 1 such that $(b_1, \dots, b_M) \leq (a_1, \dots, a_N)$. When computing $M_t(\alpha)$ we need only consider the minimal N -tuples. In our case, the minimal N -tuples are

$$(30) \quad (2, 15) \quad (3, 10) \quad (7, 5) \quad \text{and} \quad (2, 3, 7).$$

TABLE 1. Factorizations of $7/30$

Factorization of $7/30$	Corresponding (non-logarithmic) Mahler measures
$\frac{7}{30}$	(30)
$7 \cdot \frac{1}{30}$	(7,30)
$\frac{7}{2} \cdot \frac{1}{15}$	(7,15)
$\frac{1}{2} \cdot \frac{7}{15}$	(2,15)
$\frac{7}{3} \cdot \frac{1}{10}$	(7,10)
$\frac{1}{3} \cdot \frac{7}{10}$	(3,10)
$\frac{7}{6} \cdot \frac{1}{5}$	(7,5)
$\frac{1}{6} \cdot \frac{7}{5}$	(6,7)
$\frac{7}{2} \cdot \frac{1}{3} \cdot \frac{1}{5}$	(7,3,5)
$\frac{1}{2} \cdot \frac{7}{3} \cdot \frac{1}{5}$	(2,7,5)
$\frac{7}{2} \cdot \frac{1}{3} \cdot \frac{1}{5}$	(2,3,7)
$\frac{1}{2} \cdot \frac{7}{15} \cdot 7$	(2,15,7)
$\frac{7}{3} \cdot \frac{1}{10} \cdot 7$	(3,10,7)
$\frac{1}{3} \cdot \frac{7}{5} \cdot 7$	(6,5,7)
$\frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{5} \cdot 7$	(2,3,5,7)

Therefore, it makes sense to define the functions

$$\begin{aligned}
 f_1(t) &= \log 30 \\
 f_2(t) &= ((\log 2)^t + (\log 15)^t)^{1/t} \\
 f_3(t) &= ((\log 3)^t + (\log 10)^t)^{1/t} \\
 f_4(t) &= ((\log 7)^t + (\log 5)^t)^{1/t} \\
 f_5(t) &= ((\log 2)^t + (\log 3)^t + (\log 7)^t)^{1/t}
 \end{aligned}
 \tag{2.1}$$

and note that

$$\mu_{7/30}(t) = \min\{f_n(t) : 1 \leq n \leq 5\}.
 \tag{2.2}$$

The graphs of the functions (2.1) are given in Figure 1. Note that we appear to have an exceptional point at 1 and another exceptional point t satisfying the equation

$$((\log 10)^t + (\log 3)^t)^{1/t} = ((\log 7)^t + (\log 3)^t + (\log 2)^t)^{1/t}.$$

The apparent graph of (2.2) is given in Figure 2.

3. THE TOPOLOGIES INDUCED BY THE t -METRIC MAHLER MEASURES

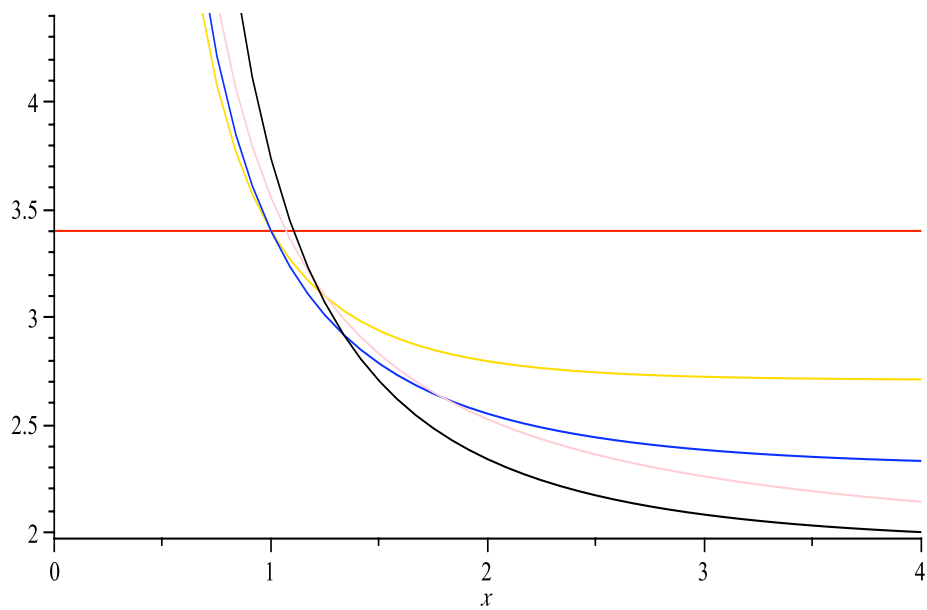
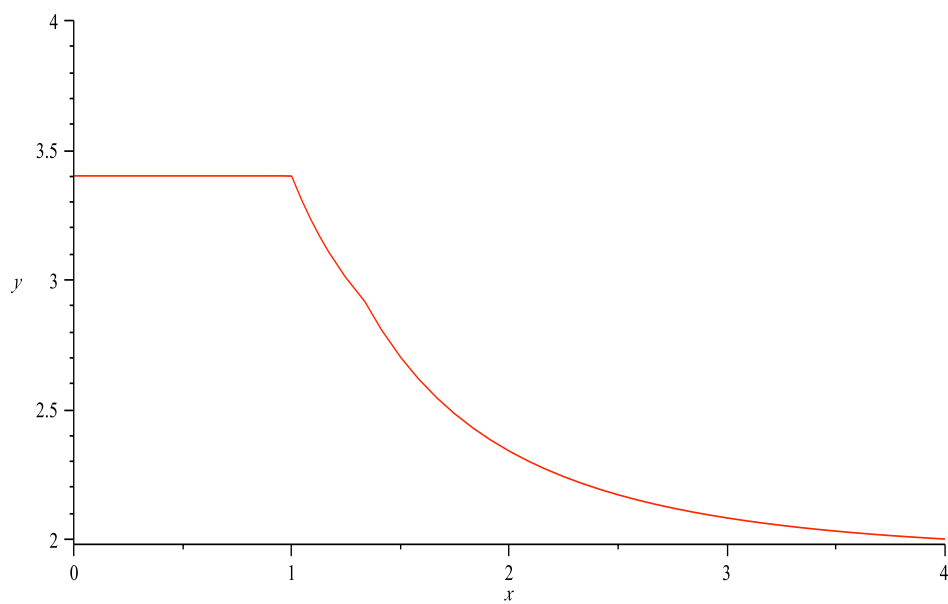
In order to proceed with the proof of Theorem 1.2, we must recall some definitions and results of [11] and [12]. If S is any subset of $\overline{\mathbb{Q}}^\times$, we write

$$\text{Rad}(S) = \left\{ \alpha \in \overline{\mathbb{Q}}^\times : \alpha^r \in S \text{ for some } r \in \mathbb{N} \right\}.$$

If K is a number field and α is an algebraic number, let K_α denote the Galois closure of $\mathbb{Q}(\alpha)$ over \mathbb{Q} . We begin with the precise statement of Lemma 3.1 of [11].

Lemma 3.1. *Let K be a Galois extension of \mathbb{Q} . If $\gamma \in \text{Rad}(K)$ then there exists a root of unity ζ and $L, S \in \mathbb{N}$ such that $\zeta\gamma^L \in K$ and*

$$M(\gamma) = S \cdot M(\zeta\gamma^L).$$

FIGURE 1. Graphs corresponding to minimal representations of $7/30$ FIGURE 2. The graph of $\mu_{7/30}(t)$ assuming Conjecture 2.1

In particular, the set

$$\{M(\gamma) : \gamma \in \text{Rad}(K), M(\gamma) \leq B\}$$

is finite for every $B \geq 0$.

It is an easy consequence of Lemma 3.1 that $M(\gamma)$ is bounded below by the Mahler measure of an element in K . Indeed, we have that

$$M(\gamma) = S \cdot M(\zeta\gamma^L) \geq M(\zeta\gamma^L)$$

and $\zeta\gamma^L \in K$. Recall that

$$C(\alpha) = \inf\{M(\gamma) : \gamma \in K_\alpha \setminus \text{Tor}(\overline{\mathbb{Q}}^\times)\}$$

and that $C(\alpha) > 0$ by Northcott's Theorem [10]. We now see easily that

$$(3.1) \quad M(\gamma) \geq C(\alpha)$$

for all $\gamma \in \text{Rad}(K_\alpha) \setminus \text{Tor}(\overline{\mathbb{Q}}^\times)$. We showed in Theorem 1.1 of [12] that the infimum in $M_t(\alpha)$ is always attained.

Theorem 3.2. *Suppose α is a non-zero algebraic number and $t \in (0, \infty]$. Then there exists a point*

$$(\alpha_1, \alpha_2, \dots) \in \tau^{-1}(\alpha) \cap \mathcal{X}(\text{Rad}(K_\alpha))$$

such that $M_t(\alpha) = \|(M(\alpha_1), M(\alpha_2), \dots)\|_t$.

Recalling that $V = \overline{\mathbb{Q}}^\times / \text{Tor}(\overline{\mathbb{Q}}^\times)$, we may proceed with our proof of Theorem 1.2.

Proof of Theorem 1.2. If Lehmer's conjecture is true, then it follows from the results of [4] that M_∞ induces the discrete topology on V . Furthermore, we always have that $M_t(\alpha) \geq M_\infty(\alpha)$ for all $\alpha \in V$, implying that M_t induces the discrete topology as well, establishing one direction of the theorem.

Now assume that Lehmer's conjecture is false and that the topologies induced by M_t and M_∞ are equivalent. Therefore, the M_t ball of radius 1 centered at 1,

$$B = \{\bar{\gamma} \in V : M_t(\bar{\gamma}) < 1\},$$

is open with respect to M_∞ . Therefore, there exists $r > 0$ such that the M_∞ -ball

$$(3.2) \quad B_0 = \{\bar{\gamma} \in V : M_\infty(\bar{\gamma}) \leq r\} \subset B.$$

We have assumed that Lehmer's conjecture is false so there exists a non-trivial point $\bar{\alpha} \in B_0$. If s is a positive integer, then the strong triangle inequality implies that $M_\infty(\bar{\alpha}^s) \leq M_\infty(\bar{\alpha}) \leq r$ so that $\bar{\alpha}^s \in B_0$ for all $s \in \mathbb{N}$. It follows from (3.2) that

$$(3.3) \quad \bar{\alpha}^s \in B$$

for all $s \in \mathbb{N}$. We will now show that $M_t(\bar{\alpha}^s)$ tends to ∞ as $s \rightarrow \infty$.

Select a point $\alpha \in \overline{\mathbb{Q}}^\times$ whose image in V equals $\bar{\alpha}$. In this case, α is not a root of unity. By Theorem 3.2, there exists a root of unity ζ and points

$$\alpha_1, \dots, \alpha_N \in \text{Rad}(K_\alpha) \setminus \text{Tor}(\overline{\mathbb{Q}}^\times)$$

such that

$$\alpha^s = \zeta \alpha_1 \cdots \alpha_N$$

and

$$(3.4) \quad M_t(\alpha^s)^t = \sum_{n=1}^N M(\alpha_n)^t.$$

Recall that the Weil height on $\alpha \in \overline{\mathbb{Q}}$ is given by

$$h(\alpha) = \frac{M(\alpha)}{\deg \alpha}.$$

Using (3.4), we have that

$$\begin{aligned} M_t(\alpha^s)^t &= \sum_{n=1}^N M(\alpha_n)^{t-1} M(\alpha_n) \\ &\geq \sum_{n=1}^N M(\alpha_n)^{t-1} h(\alpha_n) \\ &\geq \min_{1 \leq n \leq N} \{M(\alpha_n)\}^{t-1} \cdot \sum_{n=1}^N h(\alpha_n) \end{aligned}$$

It is well-known that the Weil height has the triangle inequality $h(\alpha\beta) \leq h(\alpha) + h(\beta)$ as well as the identity $h(\alpha) = h(\zeta\alpha)$ for all roots of unity ζ . It follows that

$$\begin{aligned} M_t(\alpha^s)^t &\geq \min_{1 \leq n \leq N} \{M(\alpha_n)\}^{t-1} \cdot h(\alpha_1 \cdots \alpha_N) \\ &= \min_{1 \leq n \leq N} \{M(\alpha_n)\}^{t-1} \cdot h(\alpha^s) \end{aligned}$$

Furthermore, we have that $h(\alpha^r) = |r| \cdot h(\alpha)$ for all integers r . This leaves

$$(3.5) \quad M_t(\alpha^s)^t \geq s \cdot h(\alpha) \cdot \min_{1 \leq n \leq N} \{M(\alpha_n)\}^{t-1}$$

We know that α is not a root of unity so that $h(\alpha) > 0$. Also, We know that $\alpha_n \in \text{Rad}(K_\alpha) \setminus \text{Tor}(\overline{\mathbb{Q}}^\times)$ for all n . It follows from (3.1) that $M(\alpha_n) \geq C(\alpha)$ for all n . By (3.5), we obtain that

$$M_t(\alpha^s)^t \geq s \cdot h(\alpha) \cdot C(\alpha)^{t-1},$$

the right hand side of which tends to infinity as $s \rightarrow \infty$. This proves that $\bar{\alpha}^s \notin B$ for sufficiently large s , contradicting (3.3). \square

4. α -STANDARD AND α -EXCEPTIONAL POINTS

All of our proofs regarding α -standard and α -exceptional points are based upon the following result.

Theorem 4.1. *Let α be a non-zero algebraic number and T a positive real number. Then there exists a finite collection of points $\mathcal{X} = \mathcal{X}(\alpha, T) \subseteq \mathcal{X}(\mathbb{R})$ such that*

$$M_t(\alpha) = \min\{\|\mathbf{x}\|_t : \mathbf{x} \in \mathcal{X}\}$$

for all $t \leq T$.

Proof. By Lemma 3.1, the set

$$(4.1) \quad R(\alpha) = \{M(\gamma) : \gamma \in \text{Rad}(K_\alpha) \text{ and } M(\gamma) \leq M(\alpha)\}$$

is finite and $C(\alpha) = \min R(\alpha) \setminus \text{Tor}(\overline{\mathbb{Q}}^\times)$. We also note that $M(\alpha) \geq C(\alpha) > 0$. Next, we define

$$J = J(\alpha, T) = \left\lfloor \left(\frac{M(\alpha)}{C(\alpha)} \right)^T + 1 \right\rfloor.$$

Finally, we write

$$\mathcal{X} = \left\{ (M(\alpha_1), \dots, M(\alpha_N), 0, 0, \dots) : M(\alpha_n) \in R(\alpha), N \leq J(\alpha, T) \text{ and } \alpha = \prod_{n=1}^N \alpha_n \right\}.$$

We claim that \mathcal{X} is finite and that

$$(4.2) \quad M_t(\alpha) = \min\{\|\mathbf{x}\|_t : \mathbf{x} \in \mathcal{X}\}$$

for all $t \leq T$. We have immediately that \mathcal{X} injects into

$$\underbrace{R(\alpha) \times \cdots \times R(\alpha)}_{J \text{ times}}.$$

Since each set $R(\alpha)$ is finite, it follows that \mathcal{X} is finite.

Now we must verify (4.2). By the definition of $M_t(\alpha)$, we see quickly that

$$(4.3) \quad M_t(\alpha) \leq \min\{\|\mathbf{x}\|_t : \mathbf{x} \in \mathcal{X}\}.$$

To show that we always have equality in (4.3), we must show that, for every positive real $t \leq T$, there exists $\mathbf{x} \in \mathcal{X}$ such that $M_t(\alpha) = \|\mathbf{x}\|_t$. By Theorem 3.2, we know there exist points $\alpha_1, \alpha_2, \dots, \alpha_N \in \text{Rad}(K_\alpha)$ such that $\alpha = \alpha_1 \cdots \alpha_N$ and

$$(4.4) \quad M_t(\alpha) = \|(M(\alpha_1), \dots, M(\alpha_N), 0, 0, \dots)\|_t.$$

We may assume without loss of generality that at most one of $\alpha_1, \dots, \alpha_N$ is a root of unity. Now we write

$$\mathbf{m} = (M(\alpha_1), \dots, M(\alpha_N), 0, 0, \dots)$$

so we have that

$$M_t(\alpha) = \|\mathbf{m}\|_t.$$

We must show that $\mathbf{m} \in \mathcal{X}$.

By our above remarks, we know that $\alpha_n \in \text{Rad}(K_\alpha)$ for all n . Furthermore, we have that $M_t(\alpha) \leq M(\alpha)$, so we also obtain that $M(\alpha_n) \leq M(\alpha)$, which implies that $M(\alpha_n) \in R(\alpha)$. For every n such that α_n is not a root of unity, we have that $M(\alpha_n) \geq C(\alpha)$ so we obtain

$$M(\alpha)^t \geq M_t(\alpha)^t = \sum_{n=1}^N M(\alpha_n)^t \geq (N-1) \cdot C(\alpha)^t,$$

and therefore,

$$N-1 \leq \left(\frac{M(\alpha)}{C(\alpha)} \right)^t.$$

It is clear that $M(\alpha) \geq C(\alpha)$ which yields

$$(4.5) \quad N \leq J(\alpha, T)$$

showing that $\mathbf{m} \in \mathcal{X}$ and completing the proof. \square

We noted earlier that the continuity of μ_α was proved in [12]. However, Theorem 4.1 gives us a much simpler proof.

Corollary 4.2. μ_α is continuous on $(0, \infty)$.

Proof. On an interval $(0, T]$, Theorem 4.1 establishes that μ_α is the minimum of a finite number of continuous functions. It follows that μ_α is itself continuous. \square

Before we can prove Theorem 1.3, we give one additional definition along with a lemma. For a positive real number T and an algebraic number α , we will, for the remainder of this paper, let $\mathcal{X} = \mathcal{X}(\alpha, T)$ be as in the conclusion of Theorem 4.1. We say that $s \leq T$ is an *intersection point with respect to \mathcal{X}* if there exist $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ such that $\|\mathbf{x}\|_s = \|\mathbf{y}\|_s$ but $t \mapsto \|\mathbf{x}\|_t$ is not the same function as $t \mapsto \|\mathbf{y}\|_t$.

Lemma 4.3. *Suppose that α is a non-zero algebraic number and T is a positive real number. If $I \subseteq (0, T]$ is an interval containing no intersection points with respect to $\mathcal{X}(\alpha, T)$ then I is α -uniform.*

Proof. Assume that I is not α -uniform and fix a point $t \in I$. By definition of α -uniform, for every point $\mathbf{x} \in \mathcal{X}$ such that $M_t(\alpha) = \|\mathbf{x}\|_t$, there exists $s \in I$ such that $M_s(\alpha) < \|\mathbf{x}\|_s$. We may select $\mathbf{y} \in \mathcal{X}$ such that $M_s(\alpha) = \|\mathbf{y}\|_s$ and note that $M_t(\alpha) \leq \|\mathbf{y}\|_t$. Hence, we have that

$$\|\mathbf{x}\|_t \leq \|\mathbf{y}\|_t \text{ and } \|\mathbf{x}\|_s > \|\mathbf{y}\|_s.$$

By the Intermediate Value Theorem, there exists a point r between s and t such that $\|\mathbf{x}\|_r = \|\mathbf{y}\|_r$. This means that I contains an intersection point, a contradiction. \square

We are now prepared to prove Theorem 1.3.

Proof of Theorem 1.3. We first show that there are only finitely many intersection points of \mathcal{X} . Let

$$\mathbf{x} = (x_1, \dots, x_N, 0, 0, \dots) \text{ and } \mathbf{y} = (y_1, \dots, y_M, 0, 0, \dots)$$

be elements of \mathcal{X} such that $x_n, y_n \geq 0$. Further suppose that $t \mapsto \|\mathbf{x}\|_t$ and $t \mapsto \|\mathbf{y}\|_t$ are distinct functions. Now write

$$F(z) = \sum_{n=1}^N x_n^z - \sum_{m=1}^M y_m^z$$

and note that $F(z)$ is an entire function with $F \not\equiv 0$. If F has infinitely many zeros $[0, T]$, then these zeros have a cluster point in \mathbb{C} , a contradiction. So F may only have finitely many zeros in $[0, T]$, and hence, the functions $\|\mathbf{x}\|_t$ and $\|\mathbf{y}\|_t$ may only intersect in finitely many points in $[0, T]$. It now follows that there are only finitely many intersection points.

Next, assume that t is not an intersection point. Since the set of intersection points is finite, we know there exists a neighborhood I of t that contains no intersection points. It now follows from Lemma 4.3 that I is α -uniform so that t is α -standard. In other words, we have shown that every α -exceptional point in $(0, T)$ must also be an intersection point. However, there are only finitely many intersection points, so there are only finitely many α -exceptional points in $(0, T)$. \square

We now proceed with the proof of Theorem 1.4, which requires the following two lemmas. The first of these lemmas shows that even α -exceptional points have neighborhoods that are relatively well behaved.

Lemma 4.4. *If $t \in (0, \infty)$ then there exists a neighborhood (a, b) of t such that $(a, t]$ and $[t, b)$ are α -uniform.*

Proof. If t is α -standard, then the result is obvious, so we may assume that t is α -exceptional.

Set $T = t + 1$ and let $\mathcal{X} = \mathcal{X}(\alpha, T)$ be the set from the conclusion of Theorem 4.1. Since \mathcal{X} has only finitely many intersection points, there must exist a neighborhood (a, b) of t containing no intersection points except t . In particular, (t, b) contains no intersection points, so it follows from Lemma 4.3 that (t, b) is α -uniform. Therefore, there exists $\mathbf{x} \in \mathcal{X}$ such that $M_s(\alpha) = \|\mathbf{x}\|_s$ for all $s \in (t, b)$.

By Theorem 4.2, we know that μ_α is continuous on $[t, b)$. Of course, $s \mapsto \|\mathbf{x}\|_s$ is also continuous on this interval so that

$$M_t(\alpha) = \lim_{s \rightarrow t^+} M_s(\alpha) = \lim_{s \rightarrow t^+} \|\mathbf{x}\|_s = \|\mathbf{x}\|_t$$

showing that $M_s(\alpha) = \|\mathbf{x}\|_s$ for all $s \in [t, b)$. This establishes that $[t, b)$ is α -uniform. A similar argument is used to show that $(a, t]$ is α -uniform, completing the proof. \square

Our next lemma shows that, in order to prove that an interval I is α -uniform, we need only show the existence of a cover of I by α -uniform open intervals. Here, we understand that open means open with respect to S .

Lemma 4.5. *Suppose $S \subset (0, \infty)$ is any interval. If there exists a finite cover of S by α -uniform open intervals, then S is α -uniform.*

Proof. Suppose $\{I_n : 1 \leq n \leq N\}$ is a collection of open intervals in $(0, \infty)$ such that

$$S = \bigcup_{n=1}^N I_n$$

and I_n is α -uniform for all n . Since S is connected, we must have that

$$I_1 \cap \left(\bigcup_{n=2}^N I_n \right) \neq \emptyset$$

so that there exists some k such that $I_1 \cap I_k \neq \emptyset$. Since both I_1 and I_k are open intervals, their intersection must be a non-empty open interval. We know that I_1 and I_k are α -uniform, so there exist points $(x_1, \dots, x_L, 0, 0, \dots), (y_1, \dots, y_M, 0, 0, \dots) \in \mathcal{X}(\mathbb{R})$ such that

$$M_t(\alpha)^t = \sum_{l=1}^L x_l^t \quad \text{for all } t \in I_1$$

and

$$M_t(\alpha)^t = \sum_{m=1}^M y_m^t \quad \text{for all } t \in I_k.$$

These functions must be equal on the open interval $I_1 \cap I_k$. That is, we have that

$$(4.6) \quad \sum_{l=1}^L x_l^z = \sum_{m=1}^M y_m^z$$

on a set having a limit point in \mathbb{C} . Since both sides of (4.6) are entire functions, we conclude that they must be equal on all of \mathbb{C} . In particular, we have shown that

$$M_t(\alpha)^t = \sum_{l=1}^L x_l^t \quad \text{for all } t \in I_1 \cup I_k$$

implying that $I_1 \cup I_k$ is α -uniform. We now see that the set of intervals

$$\{I_1 \cup I_k\} \cup \{I_n : 2 \leq n \leq N \text{ and } n \neq k\}$$

is a cover of S by $N - 1$ α -uniform open intervals. Repeating the above argument $N - 1$ more times, we obtain a cover containing only one interval. \square

In view of the above lemmas, the proof of Theorem 1.4 is fairly straightforward.

Proof of Theorem 1.4. If $[a, b]$ is α -uniform, then it is clear that every point in (a, b) is α -standard. Similarly, if $(0, a]$ is α -uniform then every point in $(0, a)$ is α -standard. We now prove the opposite directions of both statements beginning with the first.

Assume now that every point in (a, b) is α -standard. Hence, there exists a cover \mathcal{I} of (a, b) by α -uniform open intervals. Furthermore, by Lemma 4.4, there exist points $c, d \in (a, b)$ such that the intervals

$$J_1 = [a, c) \quad \text{and} \quad J_2 = (d, b]$$

are α -uniform. Therefore, the collection of intervals

$$\{J_1\} \cup \{J_2\} \cup \mathcal{I}$$

forms a cover of $[a, b]$ by α -uniform intervals which are all open with respect to $[a, b]$. Since $[a, b]$ is compact there exists a finite subcover and the result follows from Lemma 4.5.

To prove the second statement, recall that [12] establishes μ_α to be constant in a neighborhood of 0. In particular, there exists $\varepsilon > 0$ such that $(0, 2\varepsilon)$ is α -uniform. We know that $(0, a)$ contains no α -standard points, so that (ε, a) does not either. By the first statement of this theorem, we know that $(\varepsilon, a]$ is α -uniform. Certainly

$$(0, 2\varepsilon) \cup (\varepsilon, a]$$

is a finite cover of $(0, a]$ by α -uniform intervals that are open in $(0, a]$. It follows from Lemma 4.5 that $(0, a]$ is α -uniform. \square

Equipped with Theorems 1.3 and 1.4, we can give our proof of Corollary 1.5.

Proof of Corollary 1.5. By Theorem 1.3, there are finitely many exceptional points in $(0, T)$. Suppose these points are given by

$$0 < t_1 < t_2 < \cdots < t_N < T.$$

We write $I_0 = (0, t_1]$, $I_N = [t_N, T]$ and $I_n = [t_n, t_{n+1}]$ for all other values of n . We write

$$\mathcal{I} = \bigcup_{n=0}^N \{I_n\}$$

and claim that \mathcal{I} satisfies the required properties. Clearly, \mathcal{I} is a finite set of non-overlapping closed intervals with

$$(0, T] = \bigcup_{I \in \mathcal{I}} I,$$

which establishes (ii). The interior of I_n contains only α -standard points, so by Theorem 1.4, I_n is α -uniform for all n , verifying (i).

Now assume that $t \in (0, T)$ is α -exceptional. By (i), t must lie at an endpoint of an interval $I \in \mathcal{I}$, so that t must lie at point where two intervals from \mathcal{I} intersect. If

$t \in [t_{n-1}, t_n] \cap [t_n, t_{n+1}]$, then $t = t_n$ implying that t is α -exceptional and verifying (iii). \square

Finally, we may proceed with the proof of Theorem 1.6.

Proof of Theorem 1.6. If s is α -standard, then there exists $\mathbf{x} \in \mathcal{X}$ and a neighborhood I of s such that

$$(4.7) \quad M_t(\alpha) = \|\mathbf{x}\|_t$$

for all $t \in I$. Certainly, the right hand side of (4.7) is infinitely differentiable as a function of t for all positive t .

Assume now that μ_α is infinitely differentiable at s . By Lemma 4.4, there exists a neighborhood (a, b) of s such that $(a, s]$ and $[s, b)$ are α -uniform. Suppose that $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ are such that $M_t(\alpha) = \|\mathbf{x}\|_t$ for all $t \in (a, s]$ and $M_t(\alpha) = \|\mathbf{y}\|_t$ for all $t \in [s, b)$. Now write

$$f(z) = \|\mathbf{x}\|_z^z \quad \text{and} \quad g(z) = \|\mathbf{y}\|_z^z$$

and observe that f and g are entire functions. Moreover, their Taylor series expansions at s , given by

$$(4.8) \quad f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(s)}{n!} (z-s)^n \quad \text{and} \quad g(z) = \sum_{n=0}^{\infty} \frac{g^{(n)}(s)}{n!} (z-s)^n,$$

converge in all of \mathbb{C} .

For the remainder of this proof, we will write $\ell(t) = \mu_\alpha(t)^t$. By our assumption, ℓ is infinitely differentiable at s . We also have that $\ell(t) = f(t)$ for all $t \in (a, s]$ which implies that ℓ must also be infinitely differentiable in (a, s) . It follows easily that

$$(4.9) \quad \ell^{(n)}(t) = f^{(n)}(t) \quad \text{for all } t \in (a, s).$$

We now prove by induction that $f^{(n)}(s) = \ell^{(n)}(s)$. By the definitions of our functions, we obtain immediately $f(s) = \ell(s)$ establishing the base case. Assuming now that $f^{(n)}(s) = \ell^{(n)}(s)$, we may write

$$\ell^{(n+1)}(s) = \lim_{h \rightarrow 0} \frac{\ell^{(n)}(s+h) - \ell^{(n)}(s)}{h} = \lim_{h \rightarrow 0^-} \frac{\ell^{(n)}(s+h) - f^{(n)}(s)}{h}.$$

However, using (4.9), it follows that $f^{(n)}(s+h) = \ell^{(n)}(s+h)$ for h sufficiently close to 0. We now have that

$$(4.10) \quad \ell^{(n+1)}(s) = \lim_{h \rightarrow 0^-} \frac{f^{(n)}(s+h) - f^{(n)}(s)}{h}.$$

We know already that f must be infinitely differentiable at s , so that the right hand side of (4.10) must equal $f^{(n+1)}(s)$ establishing our claim that $f^{(n)}(s) = \ell^{(n)}(s)$ for all n .

A similar argument can be used to show that $g^{(n)}(s) = \ell^{(n)}(s)$, and therefore $g^{(n)}(s) = f^{(n)}(s)$. It now follows from (4.8) that $f(z) = g(z)$ for all $z \in \mathbb{C}$. In particular, we have shown that

$$\mu_\alpha(t) = \|\mathbf{x}\|_t = \|\mathbf{y}\|_t$$

for all $t \in (a, b)$, proving that (a, b) is α -uniform. It follows that s is α -standard. \square

5. PROOF OF THEOREM 2.2

Proof of Theorem 2.2. Since we are assuming Conjecture 2.1, we have that there exist positive integers $r_1, \dots, r_N, s_1, \dots, s_N$ such that

$$(5.1) \quad \frac{r}{s} = \prod_{n=1}^N \frac{r_n}{s_n}$$

and

$$(5.2) \quad M_t \left(\frac{r}{s} \right)^t = \sum_{n=1}^N M \left(\frac{r_n}{s_n} \right)^t = \sum_{n=1}^N \max\{|r_n|, |s_n|\}^t$$

Suppose that $\gcd(r_i, s_j) > 1$ for some i and j so there exists a prime number p such that $p \mid r_i$ and $p \mid s_j$. Now define points r'_n and s'_n , for $1 \leq n \leq N$, by

$$r'_n = \begin{cases} r_n & \text{if } n \neq i \\ r_n/p & \text{if } n = i \end{cases}$$

and

$$s'_n = \begin{cases} s_n & \text{if } n \neq j \\ s_n/p & \text{if } n = j. \end{cases}$$

We note immediately that

$$\frac{r}{s} = \prod_{n=1}^N \frac{r'_n}{s'_n}$$

and

$$\max\{|r'_n|, |s'_n|\} \leq \max\{|r_n|, |s_n|\}$$

for all n . Then using (5.2), we find that

$$\begin{aligned} M_t \left(\frac{r}{s} \right)^t &\leq \sum_{n=1}^N M \left(\frac{r'_n}{s'_n} \right)^t \\ &= \sum_{n=1}^N \max\{|r'_n|, |s'_n|\}^t \\ &\leq \sum_{n=1}^N \max\{|r_n|, |s_n|\}^t \\ &= M_t \left(\frac{r}{s} \right)^t \end{aligned}$$

implying that

$$M_t \left(\frac{r}{s} \right)^t = \sum_{n=1}^N M \left(\frac{r'_n}{s'_n} \right)^t$$

Repeating this process, we can find positive integers $a_1, \dots, a_N, b_1, \dots, b_N$ such that

$$M_t \left(\frac{r}{s} \right)^t = \sum_{n=1}^N M \left(\frac{a_n}{b_n} \right)^t$$

and each pair (a_i, b_j) are relatively prime. In particular, we have that

$$(5.3) \quad \gcd \left(\prod_{n=1}^N a_n, \prod_{n=1}^N b_n \right) = 1.$$

By (5.1), we have that

$$r \prod_{n=1}^N b_n = s \prod_{n=1}^N a_n.$$

This means that $r \mid s \prod_{n=1}^N a_n$, but since $\gcd(r, s) = 1$, we have that

$$(5.4) \quad r \mid \prod_{n=1}^N a_n.$$

However, we also know that $\prod_{n=1}^N a_n \mid r \prod_{n=1}^N b_n$, so that by (5.3), we obtain

$$\prod_{n=1}^N a_n \mid r.$$

Combining this with (5.4), we find that

$$\prod_{n=1}^N a_n = r.$$

A similar argument can be used to prove that $\prod_{n=1}^N b_n = s$ which completes the proof. \square

Finally, we provide our proof of Theorem 2.3.

Proof of Theorem 2.3. First assume that $t \leq 1$. It was shown in [3] that $M_1(\alpha) = M(\alpha)$ whenever α is rational. Using the fact that μ_α is decreasing, we have that

$$M(\alpha) = M_1(\alpha) \leq M_t(\alpha) \leq M(\alpha).$$

But $M(\alpha) = \log \alpha$ so the result follows for $t \leq 1$.

Now suppose that $t > 1$. By Theorem 2.2, there exist integers k_1, \dots, k_N such that $\alpha = k_1 \cdots k_N$ and

$$(5.5) \quad M_t(\alpha)^t = \sum_{n=1}^N M(k_n)^t = \sum_{n=1}^N (\log k_n)^t.$$

We claim that each k_n must be prime. To see this, assume there exists an integer j such that k_j is not prime and write

$$k_j = ab$$

where $a, b \in \mathbb{N}$ and $a, b > 1$. It is a straightforward application of the Mean Value Theorem to show that

$$(\log k_j)^t = (\log a + \log b)^t > (\log a)^t + (\log b)^t.$$

Applying (5.5), we find that

$$(5.6) \quad M_t(\alpha)^t > (\log a)^t + (\log b)^t + \sum_{\substack{n=1 \\ n \neq j}}^N (\log k_n)^t.$$

However, we also have that

$$\alpha = ab \cdot \prod_{\substack{n=1 \\ n \neq j}}^N k_n$$

which yields immediately

$$M_t(\alpha)^t \leq M(a)^t + M(b)^t + \sum_{\substack{n=1 \\ n \neq j}}^N M(k_n)^t = (\log a)^t + (\log b)^t + \sum_{\substack{n=1 \\ n \neq j}}^N (\log k_n)^t$$

contradicting (5.6). We have now shown that each k_n must be prime completing the proof. \square

REFERENCES

- [1] P. Borwein, E. Dobrowolski and M.J. Mossinghoff, *Lehmer's problem for polynomials with odd coefficients*, Ann. of Math. (2) **166** (2007), no. 2, 347–366.
- [2] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial*, Acta Arith. **34** (1979), no. 4, 391–401.
- [3] A. Dubickas and C.J. Smyth, *On the metric Mahler measure*, J. Number Theory **86** (2001), 368–387.
- [4] P. Fili and C.L. Samuels, *On the non-Archimedean metric Mahler measure*, J. Number Theory, **129** (2009), 1698–1708.
- [5] D.H. Lehmer, *Factorization of certain cyclotomic functions*, Ann. of Math. **34** (1933), 461–479.
- [6] M.J. Mossinghoff, *Algorithms for the determination of polynomials with small Mahler measure*, Ph.D. Thesis, University of Texas at Austin, 1995.
- [7] M.J. Mossinghoff, website, *Lehmer's Problem*, <http://www.cecm.sfu.ca/~mjm/Lehmer>.
- [8] M.J. Mossinghoff, C.G. Pinner and J.D. Vaaler, *Perturbing polynomials with all their roots on the unit circle*, Math. Comp. **67** (1998), 1707–1726.
- [9] M.J. Mossinghoff, G. Rhin and Q. Wu, *Minimal Mahler measures*, Experiment. Math. **17** (2008), no. 4, 451–458.
- [10] D.G. Northcott, *An inequality on the theory of arithmetic on algebraic varieties*, Proc. Cambridge Philos. Soc., **45** (1949), 502–509.
- [11] C.L. Samuels, *The infimum in the metric Mahler measure*, Canad. Math. Bull., to appear.
- [12] C.L. Samuels, *A collection of metric Mahler measures*, J. Ramanujan Math. Soc. **25** (2010), no. 4, 433–456.
- [13] A. Schinzel, *On the product of the conjugates outside the unit circle of an algebraic number*, Acta Arith. **24** (1973), 385–399. Addendum, ibid. **26** (1975), no. 3, 329–331.
- [14] C.J. Smyth, *On the product of the conjugates outside the unit circle of an algebraic integer*, Bull. London Math. Soc. **3** (1971), 169–175.